IONOSPHERIC IMAGING FROM A LOW EARTH ORBITER TRACKING GPS

George A. Hajj, R, Ibanez-Meier, and E. R. Kursinski
Jet Propulsion Laboratory

BIOGRAPHIES

George Hajj received his Ph.D. in Physics from Rice University in 1988. He works at the Jet Propulsion Laboratory in the Earth Orbiter System group. His main line of work is using the Global Positioning System to do precise positioning and related applications. He is presently involved in the effort at JPL to develop the GPS occultation capability for remotely sensing the earth's atmosphere and ionosphere. He is also working on developing means of reducing ionospheric and tropospheric errors that will lead to millimeter geodesy using GPS.

Rodrigo Ibanez-Meier received his Ph.]). in high-energy physics from Rice Univ. in 1992, and joined JPL's Tracking and Applications Section the same year. He is now working in POD and several applications of GPS/LEO systems, including Earth orientation, crustal motion, atmospheric and ionospheric observations.

E. Robert Kursinski is a Member of the Technical Staff of the Jet Propulsion Laboratory and is involved with developing the GPS occultation capability for remotely sensing the earth's atmosphere and ionosphere. He was the project manager and system engineer for the NASA Deep Space Network for the Voyager-Neptune encounter.

ABSTRACT

'1'orographic imaging of the ionosphere is examined using singular value decomposition analysis, The interdependency of the obtainable resolution, the accuracy of the solution and the data noise is explained. A simulation is illustrated where a true 2-I) ionospheric structure is generated, and a tomographic inversion of the structure is carried out, Inclusion of data taken in an occultation geometry reveals the strength hat is added by these data. The effect of the ionosphere on a GPS signal as viewed by a user in space is examined. Under somewhat strong solar conditions, the absolute bending of the signal is of the order of 0.01 degrees for the L1 signal; the phase advance can be as large as 90 meters.

IN'1'RO1)LJ['I'1ON

When a radio wave travels through the ionosphere, it experiences a group delay or a phase advance that is proportional to the total electron content between the transmitter and the receiver, The total electron content is defined as the line integral of the ionospheric free electron density between the transmitter and the receiver. Measurements of TEC have been achieved by different techniques with different degrees of accuracy and precision; all of these techniques involve a radio beacon satellite and a receiver. With an array of receivers tracking the same or different satellites, it is possible to obtain a set of intersecting links with a TEC measurement for each link. In such case we have a situation similar to that of X-ray projections in medical imaging where the image can be obtained using inversion methods. Unlike medical imaging however, in ionospheric imaging the geometry of the intersecting links can be quite arbitrary and therefore in many situations it lacks the sufficient strength to provide a unique inversion. in this case one can only reconstruct a solution which contains characteristics that the data arc able to resolve,

The possibility of doing 2-1) ionospheric tomography has been considered by many authors. Work by Raymund et. al. [1990] and Yeh and Raymund [1991] showed the limitation of ionospheric tomography as obtained from ground data alone. "I'his is due to the lack of links that cross the ionosphere in a tangential geometry. These links can only be provided by a transmitter and a receiver in space viewing each other in an occultation geometry. In an radio occultation geometry, a transmitting satellite is setting or rising behind the earth's atmosphere as seen by a receiver in space. Radio occultation techniques have a

heritage of over 20 years in NASA's planetary program, but the concept is however a newcomer into the earth science arena.

The first set of earth atmospheric occultations using GPS will he available when an 800" Km altitude orbiter, GPS-MET, will be launched in early 1994. GPS-MET will have a JPL developed high precision receiver, capable of tracking 8 GPS satellites simultaneously. With an all-looking antenna this low earth orbiter (1.1X)) will provide over 200" occultations daily, evenly distributed over the globe.

In this paper, we examine ways by which a LEO tracking GPS can be used to image the ionosphere. This is done in two ways, One is to combine the LEO-GPS links with TEC data Item the ground in a tomographic approach. The other is to use the LEO-GPS occultation geometry to obtain high resolution vertical profiles of the ionosphere.

in what follows we work out a 2-D example which will serve in defining the problem and answering basic questions such as the expected resolution and accuracy as a function of the geometry of the transmitting and receiving elements as well as data noise.

PEDAGOGICAL EXAMPLE

Consider the geometry of Fig. 1 , where we divide our region of interest (a rectangle of 3000x450 km^2) into 20x20 pixels. Consider a transmitting satellite to be passing above this region, and three receivers located on the ground below this region as shown in the Fig.. The corresponding links between the transmitter and the receiver are also shown. Only measurements as low as 1 3^0 elevation from the side receivers and 200 elevation from the middle receiver are included in our data set. This is to guarantee that there is no extra contribution to the TEC from outside the region of interest. We also assume that the free electron density is negligible above this regions so that TEC measurements are due only to the electron density that lies mostly within the region of interest. This assumption is not generally valid since the ionosphere extends well above 1 ()()() Km, but this example will illustrate the dependence of the sol ution on several factors such as the geometry of the links anti the choice of a particular basis to span the solution (in this example our basis is a set of pixels). This example is valid however, if the transmitting satellite is at a height comparable to 450 km. We number our pixels from 1-400," with 1 being the pixel **in the** lower left corner as shown in the figure.

Jet the electron density function be denoted by $f(\mathbf{r})$. Our- TEC_j measurement is then given by

$$TEC_j = \int_{RAYPATH\,j} f(\mathbf{r})\, ds\,; \quad j = 1, no.\ of\ links \qquad (1)$$

where the integral is along a specific link j. We choose the following basis to span our solution

$$f_i(\mathbf{r}) = 1 \text{ ifr } \in \text{ pixel } i; \quad = 0 \text{ otherwise} \qquad (2)$$

We approximate $f(\mathbf{r})$ by its digitized function, $f_d(\mathbf{r})$, which is defined by

$$f_d(\mathbf{r}) = \sum_{i=1}^{400} x_i \ f_i(\mathbf{r}), \qquad (3)$$

where x_i is the average of $f(\mathbf{r})$ in the i^{th} pixel. Our set of equat ions (Eq. (1)) can then be written in a matrix form as

$$\underset{m\times 1}{TEC} = \underset{m\times n}{D} \ \underset{n\times 1}{X} + \underset{m\times 1}{E} \ , \qquad (4)$$

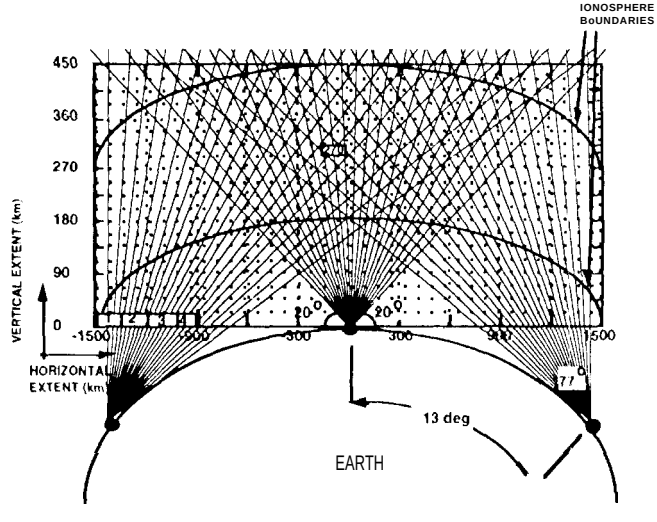

Fig. 1. Viewing geometry (off scale). Dividing the ionosphere into 20x20 pixels.

where TEC is a column of m measurements, X is a column of n unknowns, E is a column representing, the error due to data noise and discretization. D is a $m \times n$ matrix with D_{ji} being the length of link j that lies in pixel i. Each element of E, denoted by σ_j, is a function of two

errors, the noise in measurement j, and the difference between the actual measurement and the one obtained from the discrete image. Given the high precision measurements of TEC that can be obtained with the GPS system, σ_j is dominated by the latter of these two errors. Furthermore, the error introduced by discretization is also a function of (he. size of the pixel and the variability of the medium itself'. Therefore σ_j will be larger for bigger pixels and for more variability in the medium. Below we assume $\sigma_j = \sigma$ for all j's. The import ance of deter mining σ wilt become evident below.

It is important to notice that in the example of Fig. I, some pixels do not have any links passing through them, These elements in X and their corresponding null columns in D have to be removed from Eq. (4) at the outset before any inversion process is attempted. Once we do that for the example of' Fig. 1, we end up with n=392 unknown parameters. On the other hand we have m=476 measurements.

Generally, D will not be of full rank; namely, only $L < m$ equations exist that are linearly independent, Moreover, the number of independent equations $L < n$, the number of unknowns. Therefore, we have an underdetermined problem, Singular value decomposition (SVD) of the matrix D allows us to remove the linearly dependent equations from our set of equations. Performing the SVD we write

$$\underset{m \times n}{D} = \underset{m \times L}{U}\ \underset{L \times L}{W}\ \underset{L \times n}{V^T} \qquad (5)$$

where W is a diagonal $L \times L$ matrix with its diagonal elements equal to $(w_1, w_2, \ldots, w_L)$ which correspond to the non-zero singular eigen-values of D. The columns of U and V are the corresponding left and right singular eigen-vectors respectively. This factorization allows us to break the parameter space of X into two vector spaces. The first is the one spanned by the columns of $V = (v_1, v_2, \ldots, v_L)$ Each of these vectors corresponds to a linear combination of the original parameter space which can be solved for, with a variance given by

$$\sigma^{*2}_i = \frac{\sigma^2}{w_i^2} \qquad (6)$$

The second region corresponds to the null space which is the pat-I of the parameter space to which our data are not sensitive.

The singular values (SV) of our example is shown in Fig. 2 after being sorted in decreasing order. The following observations are to be made: The first 340 singular values (SV) decline by three orders of magnitudes. A sharp cutoff appears beyond the 340th singular value. The largest singular value w_1 corresponds to the best determined linear combination of the parameter space, namely the product $v_1^T \cdot X$.. The second singular value w_2 corresponds to the second best determined linear combination of the parameter space, and so on. The geometrical interpretation of' the eigenvectors v_i is that the first vector v_1 is the vector that is most nearly parallel to the rows of D; therefore, it is the most nearly parallel to the direction of our links. The second vector is the one that is most nearly parallel to the remainder, etc. [Wiggins 1972]

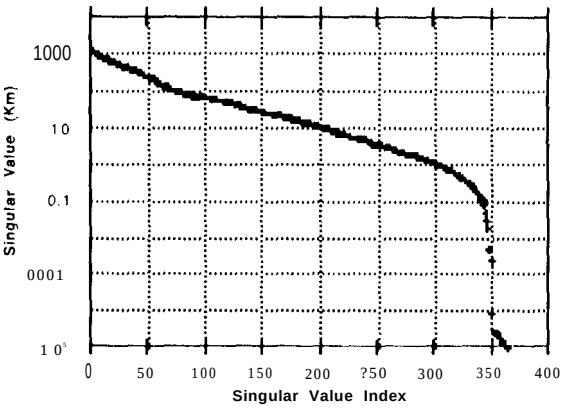

Fig. 2. Singular Values of the example of Fig. 1.

The least square solution for the system is given by

$$X = V\ W^{-1}\ U^T (TEC) \qquad (7)$$

and the solution covariance by

$$cov\,(X) = \sigma^2\, V\, W^{-2}\, V^T \qquad (8)$$

The solution given in Eq. (7) is undetermined up to an arbitrary linear combination of the vectors spanning the null space. By not including any of these vectors, we are choosing the least norm solution among ail other possible solutions. In the presence of some apriori knowledge about the ionosphere, one would set up the problem in exactly the same way as above. X would correspond to the deviation from the apriori assumed values for the electron density in each pixel, and TEC would be the difference between the measured value and the one calculated based on the apriori.

A fundamentalissue remains which is choosing the number of non-zero singular values, w_i, to include in our solution. It is important to note Item Eq. (7) that the more singular values we include in the solution, the closer we get to the true solution. on the other hand, including more and therefore smaller singular values causes the solution covariance to grow very rapidly as can be seen from Eq. (8). Different methods have been applied in choosing a cutoff value for the number Of" singular values to include. [Wiggins1972 and references therein], The main point behind most of these methods is to choose a cutoff, I_C, *such* that the variance associated with the corresponding eigenvector as given by Eq. (6) makes physical sense. For instance, in the example above, suppose we have the error in the *TEC* measurement $\sigma = 10^{16}$ (e/m^2). This error is due to data noise and discretization as explained earlier. Then a cutoff at I ()() km would corresponds to 10^{11}(e/m^3) 01 uncertainty in the electron density according to I Eq. (6). Given that the electron density to be recovered can be of the same order of magnitude Or smaller, [his implies that one should not try to include more singular values in the solution.

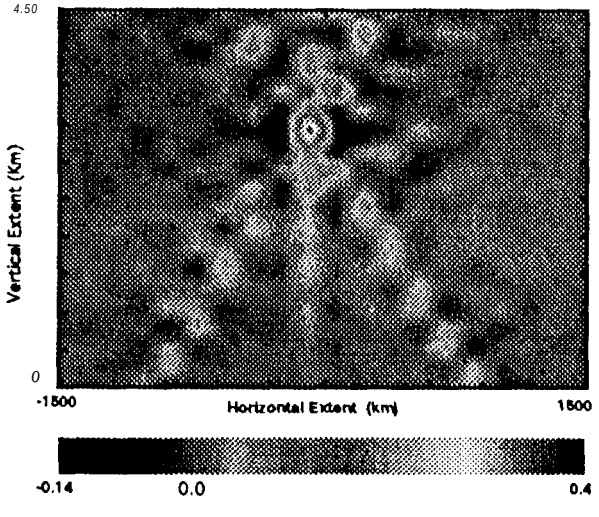

Fig.3.Thereconstruction of aunitimpulseresponse at pi xel 270, *including 100 singular values.*

Fig. 3 shows the solution for a unit impulse function in pixel 270 (see Fig. 1) which corresponds to the 270t^h column of VV^T with 100" singular values included. The reconstructed image resembles a 2-I) sine function with a peak value of .38 and a width comparable to the dimensions of the pixel in the horizontal and vertical directions.

In order to characterize the resolution of a solution, we define the resolution Of pixel i in a specific dimens On as the length of pixel i in that dimension over the height of the reconstructed impulse function of that pixel. With this definition we proceed to examine the solution in a given pixel but with the different numbers of SV asspecified above. Fig. 4 shows the horizontal and vertical resolution versus the standard deviation for pixel ?70, The number of SV included is indicated on the figure at different points. Based on this figure, if $\sigma = 10^{16}$ e/m^2, then the uncertainty in electron density al pixel 270 is about 10^{11} c/m^3 for $I_C \approx$ I 50. The corresponding resolution is about 300 km in the horizontal and 45 km in the verti cal dimensions.

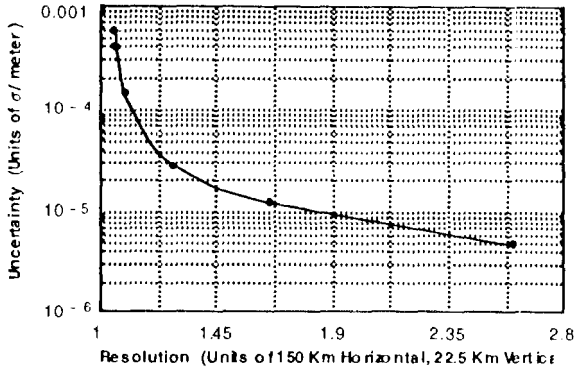

Fig. 4. Resolution and the corresponding error in solution for a given number of singular values.

SIMULATION

We proceed by generating data based on a 2-1) slice of the ion osphere (Fig. 5). This 2-I) grid which was generated by the Parameterized Ionospheric Model (PIM) corresponds to the electron density at () geomagnetic longitude, 12:00 hr. universal time, September 26, 1992, and extends form -80 to +80 latitu de. The density is representative 01 the ionosphere during the day time, with a solar spot near solar maximum. PIM is a worldwide electron density profile model, developed b y D. N. Anderson of the Air Force Phillips Laboratory. This model is a parameterized version of the Utah State i-University model, which is a complete, first principles, ionospheric model using as inputs the solar UV flux wavelength profile, the Hedin neutral wind model, the MSIS-86 neutral atmosphere model, and an electric field model obtained from all available experimental data. The 2-D profile has a bulge near 20 deg. north latitude. A vertical profile at that latitude is given in Fig. 8. The two peaks correspond to the F2 peak at 400" km altitude and the E peak at -1oo" km altitude.

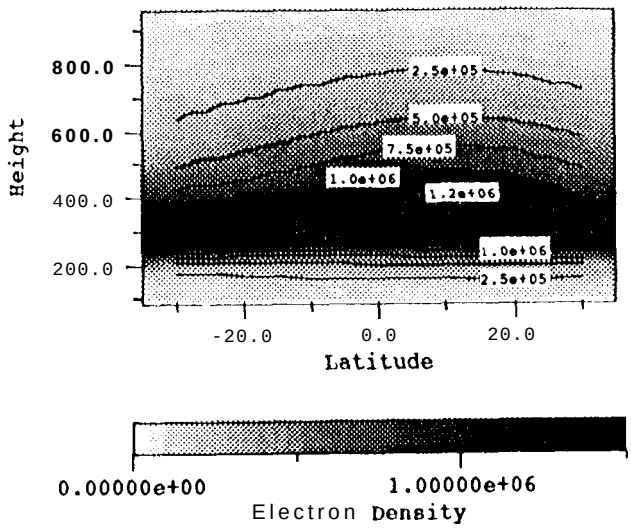

Fig. 5. Simulated electron density in e/cm^3 using the PIM model

We limit our imaging to the region between -3010 30 degree latitude and between 80 to 1000 km altitude, We divide the region into 15 *pixels* horizontality and 25 *pixels* vertically. Wc place three ground receivers at -30, 0 and 30 latitude with a data point taken every 2 degrees down to 15 degree elevation. Fig. 6.a strews the reconstruction Of the region assuming nrr apriori knowledge. Fig. 6.b shows the reconstruction using *the same* configuration as before but with an apriori electron density. The apriori used is a profile density that is uniform along the horizontal dimension. This was obtained by assuming the presence of a satellite that samples tbc region from space in an Occulting geometry. in solving for the profile, spherical symmetry was assumed. The difference between the two reconstructed images in Figs. 6.a and 6.b reflects the importance of data from space; a fact that was reflected clearly in the ionospheric tomography literature [Raymund 1992 anti references therein]. The lack of the ground data alone to capture any horizontal structure is duc to the strong correlation of the ground data at all elevations with that of the zenith measurement. This correlation makes the D matrix highly singular, and as explained earlier, only linear combinations of electron densities that are mostly parallel to the data links are well determined. Fig. 6.c is the case where the same data from ground is used in conjunction with three occultations happening in the same 2-1) plane with the tangent point of the occultations at -IS, () and 15 degree latitude and all the data are solved for sim ultaneously in a tomographic inversion. In this case, we start to distinguish some of the more detailed features of the 2-D structure, such as the presence of a bulge near 20 deg. latitude. In all of these three examples, the smallest cutoff included in the solution was equal to 1 (K) km.

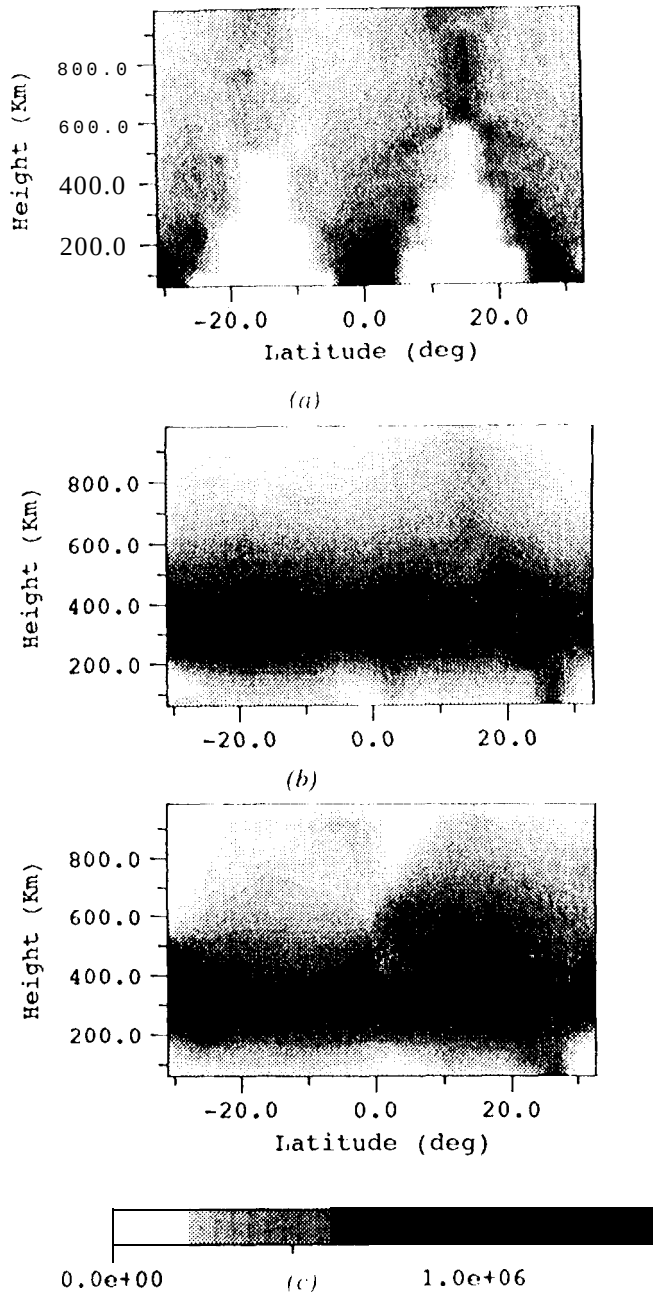

Fig. 6. Electron density (e/m^3). Tomographic reconstruction of the 2-D grid of Fig. 5. (a) without apriori assumption of a vertical structure; (b) with apriori assumption of a vertical structure; (c) combination of ground and flight data.

IONOSPHERIC OCCULTATIONS

So far flight data has been used together with ground data mostly in the context Of tomography where both horizontal and vertical resolution are being solved for simultaneously. A second manner by which the flight data can he analyzed is by a direct inversion process which gives high resolution vertical profiles of `electron density. When the signal is occulted by the atmosphere, it is Doppler shifted and bent in a way that is directly related to the index 01' refraction 01 the media. The analysis requires computing the asymptote miss distance (a) as a function of the bending angle (α) (Fig. 7); the refractive index distribution is found from $a(\alpha)$ *by* solving an Abelian integral equation under conditions of local spherical symmetry [Fjeldbo 1971]. More generally, the ionosphere can be divided into layers Of constant refractivity of predetermined arbitrary shapes. The occulting signal can then be raytraced to determine the refractivity in eat}) layer.

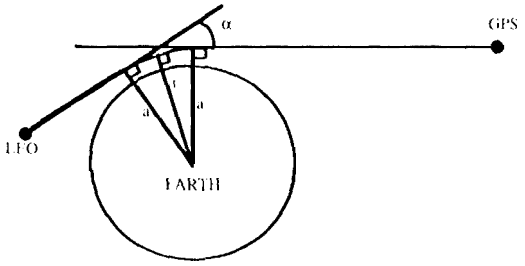

Fig. 7 The occultation geometry for the GPS-Low Earth Orbiter

The vertical resolution obtainable from an occultation measurement is set by the first Fresnel zone whit}] corresponds to about 1.5 km for a receiver in orbit at an altitude of 700 km observing a GPS satellite (Kursinski et. al., Jan. 1993). however, a true limitation in reconstructing accurate, ionospheric vertical profiles Of electron density from occultation data alone is due to the Illatively large horizontal structure that the signal has to probe causing the spherical symmetry assumption to be inaccurate. in what follows, we will examine how well one can do by ignoring the horizontal gradient, and means of improving that.

We usc Our simulated structure as before as Our truth. \Vc assume a geometry whereby a LEO is occulting a GPS satellite, with the tangent point 01 the occultation link being at 20 deg latitude (tbc tangent point is the point along the occultation link that is closest to the earth). This latitude corresponds to the region Of maximum electron density in our ?-1) grid. The profile is shown in Fig. 8.

The amount of bending and signal delay as a function of the height of the tangent point is shown in Fig. 9 for the GPS L1 signal. The negative delay implies a phase advance, A positive sign of the bending angle is such that the curvature of the ray is toward the center Of the earth. The bending of the ray is directly proportional to the amount of Doppler shift that the ray experiences, which in turn is proportional to the phase change. This explains why the bending changes sign at the point where the delay is at a minimum near the F2 peak. The oscillations in bending near the E region are similarly due to the change in curvature of the delay. It is worth noticing that even during the day time and near solar maximum, the absolute bending dots not exceed 0.015 degrees. The absolute delay On the other hand is about 90 meters, which corresponds to a total electron content of about 250×10^{16} e/m^2.

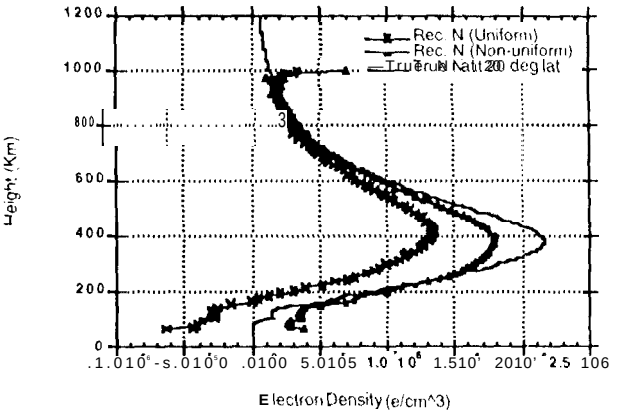

Fig. 8. True and recon structed electron density using two differen t approaches.

Given that the bending is small, in the following analysis we will assume that the signal travels along a straight path connecting the transmitter and the receiver. Assuming a locally spherically symmetric atmosphere, wc solve for the electron density profile at the tangent point, which wc choose to be at 20^{o} latitude. Fig. 8 shows the reconstructed profile (open squares, labeled by "Rec. N (Uniform)"). This solution underestimates the true profile. The reason being that the true electron density profile at 20^{o} corresponds to a maximum in the plane of occultation. This implies that the signal would be slower on both sides of the tangent point; therefore, on the average, it would give a smaller estimate of the electron density. However, under genera] conditions, when the profile corresponds to neither a maximum nor a minimum with respect to its surrounding area, a cancellation effect takes place and a more accurate reconstruction can be obtained. For instance, if for each layer that the signal is

passing through, we have a horizontal gradient that is changing linearly as a function of the arc length from the tangent point Of occultation, the signal slows down and speeds up around the tangent point by exactly the same amount, so that the estimated refractivity is exactly that of the tangent point. in the more general case, all odd components of the horizontal gradient around the tangent point cancel; the only terms that contribute to the error in refractivity are the even terms. This is explained further by Kursinski et. al (Apr. 1993). Another problem of the reconstructed profile is that it depicts a negative electron density at (he bottom of the ionosphere, which is a physical impossibility. This can be used as an extra constraint to improve the results, particularly in (he E region.

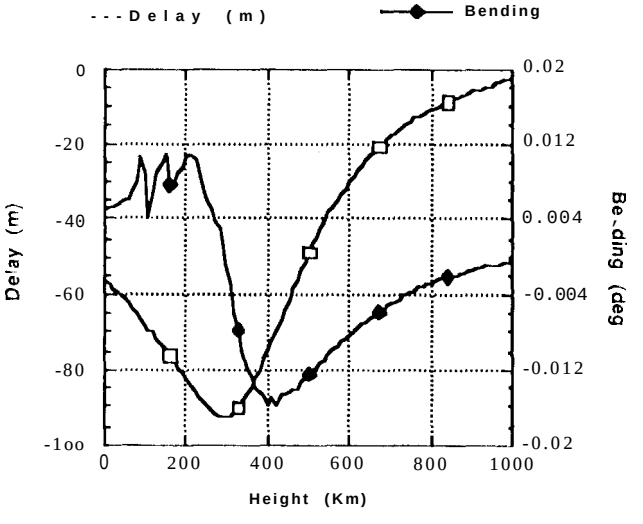

Fig. 9. Delay and bending Of the GPS L1 signal as it travels the earth's ionosphere.

A first order improvement over the spherically symmetric assumption is to divide the ionosphere into pixels in the same manner that has been done for the tomography. Then constrain pixels lying next to each other horizontally (i.e. in the same layer) to have a given functional form with an undetermined scale factor. This functional form can be the same for all layers, or different for different layers, Then One can solve for the unknown scale factor for each layer. This reduces a 2 or 3 dimensional problem into a one dimension problem where only the scale factor has to be found for each layer. In Fig. 8 we show the reconstructed profile when the same functional form is applied to all the layers of the ionosphere. The applied form is the one that describes horizontal changes in ground zenith TEC measurement. Such measurements are. available from models or real measurements. [Wilson et, al. 1993.] With the application of a horizontal gradient, the solution improves appreciably. A second order improvement is yet possible by applying different horizontal gradients to the different layers, as can be obtained from nearby occultations. "[his has been suggested and is currently under investigations by Yam Chiu and his colleagues at Lockheed.

CONCLUSION

When tomography is employed to reconstruct the ionosphere, one should always expect insufficient strength in the data to obtain a unique solution. Resolution and accuracy are very much dependent On the model used, the variability of the medium and on the data noise. Careful examination of the number of singular values to include in the solution is needed in order to obtain a sensible solution.

ionospheric tomography is greatly improved by including data obtained from space. Ground TEC measurements lack the sufficient strength to sense the vertical structure in the ionosphere. A low earth orbiter observing the GPS provides a dense global coverage of occultations which can be used by themselves or in conjunction with available ground data to obtain a global high-resolution global 3-D distribution of electrons.

ACKNOWLEDGMENTS

The work described in this paper was carried out by the Jet Propulsion laboratory, California institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

Fjeldbo, G. et. al. 1971, "The neutral atmosphere of Venus as studied with the Mariner V radio occultation experiments, " *Astronom. J.*, 76, pp. 123-140

Kursinski E. R. et. al. Jan, 1993, "Temperature or moisture profiles from radio occultation measurements", *Proc. of the 8th Symp. on Meteorological Observations and Instrumentation*, pp. J153-J 158, Amer. Meteorol Sot., Anaheim, CA.

Kursinski E.R. et. al. Apr. 1993, "Atmospheric profiles from radio occultation measurements Of GPS satellites", *Proc. of the International Soc. for Optical Engineering symp. on Optical Engineering and Photonics in Aerospace Science and Sensing*, Orlando, FL. Paper no. 1935-13.

Raymund T.D. et. al. 1992, "Ionospheric tomography: Its limitations and reconstruction methods", submitted to *JATP* April 28, 1992.

RaymundT. D. et. **al. 1990,** "Applicationof computerized tomography to the investigation of ionospheric structures", *Radio Science*, V.25, No.5, *pp. 77* 1-789, Sep./Oct.

Wiggins R. A. 1972, "The general linear inverse problem: implication of surface waves and free oscillations for earth structure", *Rev. of Geophy. and Space Phys.*, V{), lo, No. 1, pp. 251-2S5.

Wilson, B. D. at. al., 1993. "Sub-daily northern hemisphere ionospheric maps using the IGS GPS network," *Proc. Of the 7t^{h} Intern. Ionospheric Effects Symp.*, J. Goodman, ed., Alexandria, VA May.

Yeh K C. and T. D. Raymund (1991), "1 imitations of ionospheric imaging by tomography", *Radio Science*, Vo. **26,** No. 6, **pp. 1361-1380,** Nov./l))ec.